

(11) Publication number:

0 223 618
A2

(12)

EUROPEAN PATENT APPLICATION

(21) Application number: 88401590.4

(51) Int. Cl.: **C 12 Q 1/68**
// G01N33/48, G01N21/33

(22) Date of filing: 18.07.86

(30) Priority: 18.07.85 US 756129

(43) Date of publication of application:
27.05.87 Bulletin 87/22

(64) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(71) Applicant: **NEW YORK MEDICAL COLLEGE**
Elmwood Hall
Valhalla New York (US)

(72) Inventor: **Melamede, Robert Jean**
Dean road
Carnel New York 10512 (US)

(74) Representative: **Warcoln, Jacques et al**
Cabinet Régimbeau 28, avenue Kléber
F-75116 Paris (FR)

The title of the invention has been amended (Guidelines for Examination in the EPO, A-III, 7.3).

(84) Automatable process for nucleotide sequencing.

(57) This invention relates to a method for determining the nucleotide sequence of DNA and RNA molecules. The method is automatable and avoids the use of radioactive labels and gel electrophoresis. The method is also adaptable for introducing site-specific mutations in DNA and RNA molecules.

EP 0 223 618 A2

Description

AUTOMATABLE PROCESS FOR SEQUENCING NUCLEOTIDE

1. FIELD OF THE INVENTION

This invention relates to a new method for determining the base sequence of either DNA or RNA. The method of the invention is automatable and does not require the use of radioactive labels. The method may also be used to alter the sequence of a DNA or RNA molecule at a specific site, thus providing for site-specific mutagenesis of the molecule.

2. BACKGROUND OF THE INVENTION

DNA is a long threadlike macromolecule comprising a chain of deoxyribonucleotides. Similarly, RNA is composed of a chain of ribonucleotides. A nucleotide consists of a nucleoside, i.e., a nitrogenous base linked to a pentose sugar, and one or more phosphate groups which is usually esterified at the hydroxyl group attached to C-5 of the pentose sugar (indicated as 5') of the nucleoside. Such compounds are called nucleoside 5'-phosphates or 5'-nucleotides. In a molecule of DNA the pentose sugar is deoxyribose, whereas in a molecule of RNA the pentose sugar is ribose. The nitrogenous base can be a purine derivative such as adenine or guanine, or a pyrimidine derivative such as cytosine, thymine (in deoxyribonucleotides) or uracil (in ribonucleotides). Thus, the major nucleotides of DNA are deoxyadenosine 5'-triphosphate (dATP), deoxyguanosine 5'-triphosphate (dGTP), deoxycytidine 5'-triphosphate (dCTP), and deoxythymidine 5'-triphosphate (dTTP). The major nucleotides of RNA are adenosine 5'-triphosphate (ATP), guanosine 5'-triphosphate (GTP), cytidine 5'-triphosphate (CTP) and uridine 5'-triphosphate (UTP).

The sequence of the purine and pyrimidine bases of the DNA or RNA molecule encodes the genetic information contained in the molecule. The sugar and phosphate groups of a DNA or RNA molecule perform a structural role, forming the backbone of the macromolecule. Specifically, the sugar moiety of each nucleotide is linked by a phosphodiester bridge to the sugar moiety of the adjacent nucleotide as follows: the 3'-hydroxyl of the pentose of one nucleotide is joined to the 5'-hydroxyl of the pentose of the adjacent nucleotide by a phosphodiester bond. One terminus of the nucleotide chain has a 5'-hydroxyl group and the other terminus of the nucleotide chain has a 3'-hydroxyl group; thus the nucleotide chain has a polarity. By convention, the base sequence of nucleotide chains are written in a 5' to 3' direction.

The formation of the phosphodiester bonds between deoxyribonucleotides is catalyzed by the enzyme DNA polymerase. DNA polymerase requires the following components to catalyze the synthesis of a chain of DNA: a template strand (e.g. a single-stranded DNA molecule), a primer (i.e., a short DNA or RNA chain with free 3'-hydroxyl group, that is hybridized to a specific site on the single-stranded template), and activated deoxyribonucleo-

tide precursors (i.e., nucleoside 5'-triphosphates or dNTPs). Elongation of the primer strand, catalyzed by DNA polymerase, proceeds in the 5' to 3' direction along the template. The reaction occurs by means of nucleophilic attack of the 3'-hydroxyl terminus of the primer on the innermost phosphorous atom of the incoming nucleotides; a phosphodiester bridge is formed and pyrophosphate is released. DNA polymerase catalyzes the formation of a phosphodiester bond only if the base of the incoming nucleotide is complementary to the base of the nucleotide on the template strand; that is, the incoming nucleotide must form the correct Watson-Crick type of base-pair with the template. Thus, DNA polymerase is a template-directed enzyme. Reverse transcriptase is also a template-directed DNA polymerase, but requires RNA as its template. Another enzyme, RNA polymerase, catalyzes the polymerization of activated ribonucleotide precursors that are complementary to the DNA template. Some polymerases, such as *E. coli* DNA polymerase I and T4 DNA polymerase, also have 3' to 5' exonuclease activity that acts on unpaired termini. This 3' to 5' exonuclease activity serves a "proof-reading" function by removing mispaired bases before polymerization continues; i.e., the mispaired bases are edited out of the elongating strand.

2.1. DNA SEQUENCING

A number of different procedures are currently used to determine the base sequence of DNA or RNA molecules. While the approaches vary considerably, every one of the methods currently used has the following common elements: (a) a method for producing a population of radioactive polynucleotides in which each molecule has one common terminus (either 5' or 3');

(b) a method for producing from this population of radioactive polynucleotides an array of polynucleotides with one common terminus but varying in length at the other terminus in increments of a single base; and

(c) a method for ordering the population of molecules by size, usually by electrophoretic separation in a high-resolution denaturing polyacrylamide gel from which an autoradiograph is prepared. The sequence is deduced from the resulting "bands" or "ladders" on the autoradiogram.

Specific sequencing methods are discussed in the subsections that follow.

2.1.1. PLUS/MINUS DNA SEQUENCING METHOD

The plus/minus DNA sequencing method (Sanger and Coulson, 1975, *J. Mol. Biol.* 94: 441-448) involves the following:

DNA polymerase is first used to elongate a primer oligonucleotide and copy the template in the presence of the four activated nucleotide precursors, one of which is labeled with ³²P. Ideally, the synthesis is non-synchronous and as random as

possible, so that the maximum number of oligonucleotides of different length, all starting from the primer, are formed. Excess unreacted nucleotides are removed and the mixture of DNA strands is divided in two. One half is treated according to the "Minus" System and the other according to the "Plus" System, as described below:

(a) The "Minus" System: The mixture of random length radiolabeled oligonucleotides, which are still hybridized to the template DNA, is divided into four separate reaction mixtures and reincubated with DNA polymerase in the presence of three activated nucleotide precursors; that is one of the four nucleoside 5'-triphosphates is missing from each reaction mixture. Elongation of each chain will proceed as far as it can along the template; in other words, each chain will terminate at its 3' end at a position before the site of incorporation of the missing residue. For example, in the -A system, dATP is the nucleotide missing from the reaction mixture and each chain will terminate at its 3' end at a position before a dATP residue would be incorporated into the growing chain. Therefore, at the end of the incubation period each reaction mixture will contain a population of DNA molecules each having a common 5' terminus but varying in length at the 3' terminus. The radiolabeled nucleotides of varying lengths in each reaction mixture are fractionated according to size by electrophoresis in a denaturing polyacrylamide gel; each reaction mixture is fractionated in a separate lane. The relative position of each residue along the DNA may be located and the sequence of DNA may be deduced from the autoradiograph of the resulting gel. This system alone is usually not sufficient to establish a sequence, so a second similar system, the Plus System, described below, is normally used in conjunction with it.

(b) The "Plus" System: The mixture of random length radiolabeled oligonucleotides, which are still hybridized to the template DNA is divided into four separate reaction mixtures each of which is reincubated with DNA polymerase in the presence of only one of the four activated nucleotide precursors. For example, in the +A system only dATP is present in the reaction mixture. While the population of DNA molecules each has a common 5' terminus, all the chains will have varying lengths that terminate with deoxyadenosine residues. The positions of the dATP residues will be indicated by bands on the autoradiograph obtained after fractionating the DNA chains in each reaction mixture according to size by electrophoresis in a denaturing polyacrylamide gel; each reaction mixture is fractionated in a separate lane. Usually these will be oligonucleotide products that are one residue larger than the corresponding bands in the -A system, but if there is more than one consecutive dATP residue, the distance between the bands in the -A and +A systems will indicate the number of such consecutive residues.

In order for the plus/minus system to yield reliable results various criteria must be satisfied. For instance, all DNA fragments must have the same 5' terminus and the Klenow fragment of DNA polymerase must be used: in order to eliminate the 5' exonuclease activity of DNA polymerase. Furthermore, it is essential that the nucleotides are fractionated according to size. Ideally, oligonucleotides of all possible lengths should be present in the initial reaction mixture so that all residues are represented in the plus and minus systems, however, it is difficult to achieve this because certain products are formed in relatively high yield whereas others are absent. It has been suggested that the polymerase acts at different rates at different sites or that this effect is partly related to the secondary structure of the template. While Sanger et al., *supra*, report that best results are obtained if synthesis is carried out for short times with a relatively high concentration of polymerase, frequently some expected products are missing. This constitutes a limitation of the method and is one reason why it is necessary to use both the plus and minus systems. Consecutive runs of a given nucleotide present the main difficulty when using the plus/minus method of sequencing.

2.1.2. DIDEOXY CHAIN TERMINATION METHOD

The "dideoxy" chain termination DNA sequencing method of Sanger et al. 1977, *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463, also makes use of the ability of DNA polymerase to synthesize a complementary radiolabeled copy of a single stranded DNA template hybridized to a short DNA primer. The synthesis is carried out in the presence of all four deoxynucleoside 5'-triphosphates (dNTPs), one or more of which is labeled with the ^{32}P , and a 2',3'-dideoxynucleoside triphosphate analog of one of the four dNTPs. Four separate incubation mixtures are prepared each of which contains only one of the four dideoxynucleotide analogs. Once the analog is incorporated, the 3' end of the growing chain is no longer a substrate for DNA polymerase and thus cannot be elongated any further. At the end of the incubation period each reaction mixture will contain a population of DNA molecules having a common 5' terminus but varying in length to a nucleotide base-specific 3' terminus. Each population of DNA molecules is then denatured and fractionated according to size by gel electrophoresis; each reaction mixture is fractionated in a separate lane. Autoradiography of the gel allows the sequence to be deduced.

The use of the single-stranded bacteriophage M13 to obtain multiple copies of the DNA sequence of interest and its "universal primer" sequence has greatly enhanced the usefulness of the dideoxy chain termination DNA sequencing method. However, the method absolutely requires fractionation of the DNA products by size and thus involves gel electrophoresis.

2.1.3. MAXAM-GILBERT METHOD

The Maxam-Gilbert method of DNA sequencing is a chemical sequencing procedure (Maxam and Gilbert, 1977, *Proc. Natl. Acad. Sci. USA* 74: 560).

After radioactively labeling either the 3' or the 5' terminus of a discrete DNA fragment, aliquots of the DNA are placed in four separate reaction mixtures, each of which partially cleaves the DNA in a base-specific manner. The resulting population of DNA in each reaction mixture is then denatured and fractionated according to size by gel electrophoresis; each reaction mixture is fractionated in a separate lane. The DNA sequence is deduced from the ladders which appear on the resulting autoradiogram.

2.1.4. RNA SEQUENCE DETERMINATION

The major RNA sequence analysis strategies employ 5' and 3' terminal labeling protocols (England and Uhlenbeck, 1978, *Nature* 275: 561). Subsequent to radiolabeling, the RNA molecules are fragmented using base-specific RNases or chemicals. Similarly to the DNA sequencing methods, each population of oligoribonucleotides is fractionated by size via high resolution electrophoresis in a denaturing polyacrylamide gel. The sequence is then deduced from the autoradiogram corresponding to the gel.

2.1.5. AUTOMATION OF SEQUENCING

Some major drawbacks to the sequencing methods described above are that they are labor intensive, time consuming and not readily automated. Current attempts at automation involve the use of densitometers to "read" the optical density of the bands or ladders on the autoradiographs. These techniques require that the gel lanes be straight and also require careful monitoring by the operator.

2.2. SITE SPECIFIC MUTAGENESIS

Methods currently used to mutagenize DNA include *in vivo* techniques which involve treatment with mutagens such as alkylating agents, mitomycin C, ionizing radiation or ultraviolet radiation, or *in vitro* techniques such as deletion loop mutagenesis induced by bisulfite. However, these methods are likely to yield multiple-base substitutions in a non-specific manner.

Several methods have been developed to generate specific base substitutions at selected sites in DNA. (For a brief review of the methods used, see Zakour et al., 1984, *Nucleic Acids Research* 12(6): 6615-6628 and Abuzua et al., 1984, *Proc. Natl. Acad. Sci.* 81: 2030-2034) One method involves inserting into a viral DNA template or recombinant DNA a synthetic oligonucleotide which encodes a pre-selected change in its nucleotide sequence. This method is efficient and can produce any type of base substitution mutation but each different mutation that is introduced requires the synthesis of a unique oligonucleotide which encodes the mutation and is complementary to the cohesive ends which must be generated on the viral or recombinant DNA.

A second method (Shortle et al., 1980, *Proc. Natl. Acad. Sci. USA* 77: 5375-5379) involves introducing a small single-strand gap in the DNA molecule followed by mis-repair DNA synthesis: i.e., the mis-incorporation of a non-complementary nucleotide in the gap. The incorporation of α -thiol nucle-

tides into the gap minimizes the excision of the non-complementary nucleotide. When deoxyribonucleoside (1-thio)-triphosphate analogs containing a sulfur atom in place of oxygen on the phosphorous are used as substrates for the synthesis of a DNA strand that is complementary to a template DNA, the analog is incorporated as a thiomonophosphate at rates similar to those of corresponding unmodified nucleoside triphosphates. However, the phosphorothioate is not hydrolyzed by the 3' to 5' exonuclease activity of either *E. coli* DNA polymerase I or T4 DNA polymerase and, therefore, the mispaired base is not edited out. Abarzua et al. (1984, *Proc. Natl. Acad. Sci.* 81: 2030-2034) report a modification of this technique using a gapped circular DNA constructed by annealing viral single-stranded circular DNA with a mixture of linear duplex DNAs that have had their 3'-hydroxyl termini processively digested with *E. coli* exonuclease III under conditions in which the resulting, newly generated 3'-hydroxyl termini present in the various hybrid molecules span the region of interest. Base changes are induced by incorporation of mis-matched 2-thiodeoxyribonucleoside triphosphate analogs, followed by DNA repair synthesis.

A third method used is based on the infidelity of certain DNA polymerases and involves the extension of a primer by a non-proofreading DNA polymerase in the presence of a single non-complementary deoxynucleotide triphosphate, after which synthesis is completed by a highly accurate DNA polymerase in the presence of all four deoxyribonucleotide substrates. In a modification of this method, Zakour et al. (1984, *Nucleic Acids Research* 12(16): 6615-6628) used T4 DNA polymerase to elongate primer termini to a position immediately adjacent to two different preselected positions on ϕ X174 templates. Then, the error-prone DNA polymerase from avian myeloblastoma virus was used to insert single non-complementary nucleotides at the designated positions with high efficiency.

3. SUMMARY OF THE INVENTION

This invention presents a new automatable method for sequencing DNA or RNA that does not require radioactivity or gel electrophoresis. The method may also be used to accomplish the site-specific mutagenesis of any DNA or RNA molecule.

The sequencing method of the present invention involves adding an activated nucleotide precursor (a nucleoside 5'-triphosphate) having a known nitrogenous base to a reaction mixture comprising a primed single-stranded nucleotide template to be sequenced and a template-directed polymerase. The reaction conditions are adjusted to allow incorporation of the nucleotides precursor only if it is complementary to the single-stranded template at the site located one nucleotide residue beyond the 3' terminus of the primer. After allowing sufficient time for the reaction to occur, the reaction mixture is washed so that unincorporated precursors are removed while the primed template and polymerase are retained in the reaction mixture. The wash or effluent is assayed for the incorporation of precursor.

sors. The methods which may be used to detect unincorporated precursors in the effluent include but are not limited to spectroscopic methods, radioactive labeling and counting, electrochemical, and conductivity methods. The detection of all of the of nucleotide precursor in the effluent that was added to the reaction mixture indicates that the added precursor was not incorporated into the growing chain and, therefore, is not part of the nucleotide sequence. If less nucleotide precursor is detected in the effluent than was added, however, this indicates that the added precursor was incorporated into the growing chain and, therefore, is the next nucleotide of the sequence.

The sequencing method of the present invention is readily automatic. For example, the reaction chamber may be attached to five reservoirs - one of each nucleotide precursor and one for a wash buffer - that feed into the chamber. The reaction chamber should also have an outlet which feeds the effluent into the detection instrument used for the assay; for example, a spectrophotometer, a scintillation counter, Geiger-Muller counter, conductivity or electrochemical cell, etc. Ideally, the assay instrument and the valves that regulate the inlet and outlet of the reaction chamber can be controlled by a computer which is programmed to select the particular nucleotide precursor to be added to the reaction mixture, to record the instrument reading of the effluent, and to determine and record which nucleotides were incorporated into the growing chain, thus ultimately providing a print-out of the nucleotide sequence.

The sequencing method of the present invention has a number of advantages over existing methods:

- (a) radioactive labels are not required (although they may be used);
- (b) fractionation of polynucleotides by size is not required;
- (c) gel electrophoresis is not required therefore, the sequence need not be deduced by reading bands or ladders on a sequencing gel; and
- (d) sequence information is acquired as the reaction proceeds, thus allowing results to be screened during sequencing.

In another embodiment of the present invention a modification of the sequencing method may be used to alter or mutagenize a DNA or RNA sequence at a particular nucleotide site within the sequence. According to this embodiment, site-specific mutagenesis is accomplished as follows: the single-step synthesis of a nucleotide strand complementary to a template strand is accomplished as described above for DNA sequencing, but the template strand has a known nucleotide sequence. Since the nucleotide sequence of the template is known, the order of the nucleotide precursors to be added step-by-step is known. The synthesis is stopped at the nucleotide residue preceding the residue which is to be altered. The next nucleotide precursor to be added to the reaction mixture is one which cannot be edited out by the polymerase under the reaction conditions in the chamber; this nucleotide base is the mutation desired in the sequence. Although the nucleotide is

mis-paired (i.e. the base is not complementary to the template strand at that residue) the nucleotide will be incorporated into the growing strand and will not be edited out by the template directed polymerase. After each desired site-specific mutation is accomplished, the synthesis of the remaining portion of the DNA or RNA molecule need not proceed in a stepwise fashion, therefore, all four activated nucleotide precursors may be added to the reaction mixture to complete the elongation.

The reaction chamber and reservoirs used in the embodiment of the invention are similar to those described for DNA sequencing above except that an extra reservoir may be required. The site-specific mutagenesis method of the present invention is automatable and can be controlled by a computer. In this case the computer is programmed to add each nucleotide of the known sequence in the proper order, to record the instrument reading of the effluent to ensure each nucleotide is incorporated, to then add the mis-matched analog, and finally to add all four nucleotides to the reaction mixture.

4. BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 is a schematic diagram of the different components (not drawn to scale) which may be used in the practice of one embodiment of the present invention.

Fig. 2 is a schematic representation of the reaction and sequence of steps which may be used for sequencing according to the present invention.

Fig. 3 is a schematic representation of the reaction and sequence of steps which may be used for site-specific mutagenesis according to the present invention.

Fig. 4 is a flow chart of a computer program that may be used according to the present invention.

Fig. 5 is a diagram of the reaction chamber used in one example of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

This invention presents a new method for sequencing DNA or RNA which is automatable and does not require the use of radioactivity or gel electrophoresis. In addition, the present invention involves a method for the site-specific mutagenesis of a DNA or RNA sequence.

5.1. SEQUENCING

The sequencing method of the invention involves the following: The single-stranded DNA or RNA molecule to be sequenced is primed at a specific site with a short oligonucleotide primer. The primed template and a template-directed polymerase are placed in a reaction chamber that allows for separation of unreacted nucleotide precursors from the primed template and polymerase. If the template is a single-stranded DNA molecule, a DNA-directed DNA or RNA polymerase may be used; if the template is a single-stranded RNA molecule, then a reverse transcriptase (i.e., an RNA-directed DNA polymerase) may be used.

One at a time, a particular activated nucleotide

precursor is added to the chamber and allowed to react. The nucleotides added may be either deoxyribonucleotides or ribonucleotides depending upon the nature of the template and polymerase used in the reaction mixture. For instance, when the template is a single-stranded DNA molecule, the polymerase used may be a DNA-directed DNA polymerase, in which case the nucleotides added should be deoxyribonucleotides. When double stranded DNA is used, the polymerase may be a DNA-directed RNA polymerase, in which case the nucleotides added should be ribonucleotides. Similarly, when the template is a single-stranded RNA molecule, the polymerase used is a reverse-transcriptase, *i.e.*, an RNA directed DNA polymerase, in which case the nucleotides added must be deoxyribonucleotides. In any case, however, only one type of nucleotide is added at a time. For example, either dATP, dTTP, dGTP or dCTP is added to the reaction but not a mixture of these deoxyribonucleotides; similarly, either ATP, UTP, GTP or CTP is added to the reaction but not a mixture of these ribonucleotides.

If the base of the added nucleotide precursor is complementary to the template, *i.e.*, if the nucleotide precursor can form a Watson-Crick type of base pair with the template at the site located one nucleotide residue beyond the 3' terminus of the primer strand, the nucleotide precursor will be incorporated into the growing chain. If the base of the nucleotide precursor is not complementary to the template strand at the site located one nucleotide residue beyond the 3' terminus of the primer strand, the nucleotide precursor will not be incorporated into the growing chain. After an adequate time is allowed for the polymerase reaction to occur the reaction chamber is "washed" in order to separate any unreacted nucleotide precursors from the primed template. Then the wash or effluent is assayed in order to determine the amount of nucleotide precursor in the effluent. The methods which may be used to detect unincorporated precursors in the effluent include but are not limited to spectroscopic methods, such as absorption or fluorescence spectroscopy; radioactive labeling and counting (provided the nucleotide precursors are radiolabeled); and electrochemical or conductivity methods.

The process described above can be automated as follows: The reaction chamber can be connected to five reservoirs which feed into the reaction chamber, one of which contains the wash buffer whereas each of the other four contain a specific nucleotide precursor. The reaction chamber should also have an outlet valve so that after the addition of a particular nucleotide precursor and allowing for an appropriate reaction time the chamber can be washed so that the effluent moves through the outlet valve into a flow-cell of the detecting instrument used in the assay. The various components of this embodiment of the invention are illustrated schematically in FIG. 1. The reservoirs and detector can be attached to a computer that records which of the nucleotide bases was added to the reaction chamber and the detector reading of the resultant effluent. Ideally, the computer can be programmed

to choose which nucleotide precursor to feed into the reaction chamber next and to record the nucleotides which were not incorporated, thus ultimately providing a printout of the sequence.

5.2. SITE-SPECIFIC MUTAGENESIS

In a second embodiment of the present invention, a modification of the sequencing method may be used as a basis for site-specific mutagenesis of any nucleotide sequence. According to this mode of the invention, a single-stranded DNA or RNA molecule having a known sequence which is to be mutagenized at a specific site is primed with a short oligonucleotide primer. The primed template and a template-directed polymerase are placed in a reaction chamber that allows for separation of unreacted nucleotide precursors from the primed template and the polymerase. If the template is a single-stranded DNA molecule, a DNA-directed DNA or RNA polymerase may be used; if the template is a single-stranded RNA molecule, then a reverse transcriptase (*i.e.*, and RNA-directed DNA polymerase) may be used.

Each activated nucleotide precursor of the known sequence is added to the chamber and allowed to react. The nucleotides added may be either deoxyribonucleotides or ribonucleotides depending upon the nature of the template and polymerase used in the reaction mixture. For instance, when the template is a single-stranded DNA molecule, the polymerase used would be a DNA-directed DNA polymerase, in which case the nucleotides added should be deoxyribonucleotides. Similarly, when the template is a single-stranded RNA molecule, the polymerase used is a reverse-transcriptase, *i.e.*, an RNA-directed DNA polymerase, in which case the nucleotides added must be deoxyribonucleotides. Alternatively, if a double-stranded DNA template is used that contains a promoter for which a specifically functional RNA polymerase is available, *i.e.*, the T7 promoter and polymerase, the nucleotides added should be ribonucleotides. In one embodiment of the present invention, only one type of nucleotide is added at a time. For example, either dATP, dTTP, dGTP or dCTP is added to the reaction but not a mixture of these deoxyribonucleotides. Similarly, either ATP, UTP, GTP or CTP is added to the reaction but not a mixture of these ribonucleotides. In a second embodiment of the method, the nucleotides may be added up to three at a time in order to speed up the synthesis. This approach is feasible because the sequence of the template is known.

After an adequate time is allowed for the polymerase reaction to occur the reaction chamber is "washed" in order to separate any unreacted nucleotide precursors from the primed template. Then the wash or effluent may be assayed in order to determine the presence or absence of the nucleotide precursor in the effluent to ensure that the nucleotide was, in fact, incorporated into the growing chain. The methods which may be used to detect unincorporated precursors in the effluent include but are not limited to spectroscopic methods, such as absorption or fluorescence spectroscopy; radioactive labeling and counting (provided

the nucleotide precursors are radiolabeled); and electrochemical or conductivity methods.

The stepwise addition and reaction of each nucleotide or groups of up to three nucleotides is continued and ultimately stopped at the nucleotide position which precedes the residue which is to be mutagenized. The next nucleotide to be added to the reaction mixture is a nucleotide that cannot be edited out of the elongating chain by the polymerase under the conditions in the reaction chamber. For example, the nucleotide may be an analog that cannot be edited out by the polymerase; this analog base is the mutation desired in the sequence. The analog is incorporated into the growing chain even though the analog base does not form a Watson-Crick type of base pair with the nucleotide residue in the template to be mutagenized; and the mis-paired analog will not be edited out of the growing strand. After an adequate time is allowed for the polymerase reaction to occur, the reaction chamber is washed and the effluent may be assayed in order to ensure the incorporation of the analog. After the desired site-specific mutation is accomplished, the synthesis of the remaining portion of the DNA or RNA molecule need not proceed in a step-wise fashion, therefore all four unmodified nucleotide precursors may be added to the reaction mixture to complete the elongation. An example of embodiment of the invention is depicted in FIG. 3.

Nucleotide analogs which may be used in the present invention include deoxyribonucleoside (1-thio)-triphosphates containing a sulfur atom in place of an oxygen atom on the phosphorus. These analogs are incorporated as thiomonophosphates at rates similar to those of corresponding unmodified nucleoside triphosphates. However, the phosphorothioate bond is not hydrolyzed by the 3' to 5' exonuclease of either *E. coli* DNA polymerase I or T4 DNA polymerase, therefore, the incorporation of the analog as a mispaired base cannot be edited out. Other analogs may be used in the practice of this embodiment of the present invention.

The process described above can be automated as follows: The reaction chamber can be connected to five reservoirs which feed into the reaction chamber, one of which contains the wash buffer whereas each of the other four contains a specific nucleotide precursor. In addition, the reaction chamber should be connected to one reservoir for each analog. The reaction chamber should also have an outlet valve so that after the addition of a particular nucleotide precursor and allowance for an appropriate reaction time, the chamber can be washed so that the effluent moves through the outlet valve into a flow-cell of the detecting instrument used in the assay. The reservoirs and assay instrument can be attached to a computer that controls the selection of the nucleotides to be added in a step-wise manner, records the successful incorporation of each nucleotide base added to the reaction chamber, and records the assay reading of the resultant effluent. Ideally, the known nucleotide sequence with the desired site-specific mutation or mutations can simply be fed into the computer, thus facilitating the process.

The subsections below describe the invention in more detail.

5.3. THE REACTION CHAMBER AND COMPONENTS

The DNA or RNA to be sequenced serves as the template in the polymerase reaction utilized in the present invention, therefore the molecule to be sequenced should be single-stranded; however, the template may be linear or circular. Thus, any single-stranded DNA or RNA molecule may be sequenced according to the method of the present invention.

Selection of the primer, polymerase and activated nucleotide precursors used in the practice of the present invention depends upon the nature of the template to be sequenced. For example, if the template to be sequenced is DNA, the primer used may be DNA, RNA or a mixture of both. The polymerase used should be a DNA-directed polymerase. If a DNA-directed DNA polymerase is used, then deoxyribonucleotide precursors will be used in the reaction; alternatively a DNA-directed RNA polymerase requires ribonucleotide precursors to be used in the reaction. However, if the strand to be sequenced is RNA, the polymerase used should be an RNA-directed DNA polymerase, in which case deoxyribonucleotide precursors will be used in the reaction.

In an other embodiment of the invention the template can be a double-stranded DNA molecule such as a chromosome that encodes a promoter. According to this mode of the invention the polymerase should be one that recognizes the promoter and initiates transcription of mRNA; therefore, the nucleotides used are ribonucleotides.

In any case, the polymerase used should have a high level of accuracy - that is the polymerase should require correct base pairing before polymerization to ensure a high level of fidelity of the reaction. Some polymerases, such as *E. coli* DNA polymerase I, have a 5' to 3' exonuclease activity; according to one mode of the present invention, a polymerase that is low in the 5' to 3' exonuclease activity is preferred. An example of a polymerase that is low in the 5' to 3' exonuclease activity is the Klenow fragment of *E. coli* DNA polymerase I. Other polymerases which may be used in the practice of the present invention include but are not limited to AMV reverse transcriptase, *E. coli* RNA polymerase, and wheat germ RNA polymerase II.

Although the volume of the reaction can vary, in the preferred embodiment of the present invention the volume should be less than one milliliter.

The reaction vessel should be constructed so that unreacted nucleotide precursors can be separated from the polymerase and reaction products. Ideally, the reaction vessel is constructed so that each of the four nucleotide precursors and one or more wash buffers feed into the reaction chamber which is provided with an outlet that feeds into the flow cell of an instrument that can be used to assay the nucleotides in the wash or effluent. The outlet is closed during the reaction time but opened at the end of the reaction to allow displacement of the

effluent when wash buffer is fed into the research chamber.

Alternatively, the reaction vessel can be part of a continuous flow system. In this embodiment, the flow rate is adjusted such that the nucleotide precursor injected upstream of the reaction vessel has sufficient time in the reaction vessel to react (if it is the next base in the sequence) with the polymerase and nucleotide template, before passing on to the detector.

If the reaction is carried out in a liquid buffer, then a membrane having an appropriate pore size interposed at the outlet valve could be used to retain the reaction products and polymerase while allowing unreacted nucleotide precursors to pass through with the effluent. In this embodiment, the pore size selected should be large enough to allow the passage of the unreacted nucleotide precursors but not the primed template of the polymerase.

In another embodiment, the polymerase or the primed template could be immobilized on an insoluble inert support, thus, the polymerization reaction will occur on the surface of the inert support. When the primed template is immobilized, the pore size of the frit interposed at the outlet valve need only be small enough to prevent passage of the support material with the DNA and primer attached. Below the outlet valve would be a molecular trap of appropriate size to retain the polymerase. By reversing the flow through this trap, the polymerase could then be passed back through the reactor with the next nucleotide precursor to be tried.

In yet another embodiment, the reaction vessel may contain a column packing material that differentially retards the movement of each component in the reaction mixture.

The porous membrane may be composed of any inert solid material, such as dialysis membrane material, nitrocellulose, or cellulose acetate, to name but a few.

5.4. SELECTION OF NUCLEOTIDE PRECURSOR TO BE ADDED TO THE REACTION MIXTURE

The polymerization reaction of the present invention should proceed synchronously so that substantially all of the primer ends have been elongated to the same position. Various precautions and approaches may be taken to ensure that each step of the reaction goes to completion and to minimize the background or noise reading obtained.

5.4.1. SEQUENCING

In the preferred embodiment of the method for sequencing DNA, a known excess of the nucleotide precursor is added to the reaction chamber, allowed to react, and then flushed past a detector. The number of sequential bases incorporated is determined according to the relative amount of nucleotide detected in the effluent. For example, if four equivalents (relative to the molar amount of DNA in the reaction chamber) of dATP was added and the integration of the effluent signal is 50% of what would be expected for that amount of dATP, it can be concluded that two sequential "A" nucleotides were incorporated. This technique serves to both

indicate repeated nucleotides in the sequence and to shorten the reaction time.

In the preferred embodiment of the method for sequencing DNA, "mistakes" or primer strands that are not synchronously elongated are carried along with each addition of a nucleotide precursor for rectification when deemed necessary because of a deteriorating signal-to-noise ratio, so as to minimize the background absorbance of the effluent. This may be accomplished by the proper selection of the appropriate nucleotide to be added to the reaction mixture for any one step. In one embodiment, in order to carry along mistakes the penultimate nucleotide that was successfully incorporated into the primer strand should be the next nucleotide precursor added to the reaction mixture. If after washing and assaying the effluent it is determined that the nucleotide precursor was not incorporated into the elongated primer strand, then each of the other three nucleotides may be tried, one at a time, in any order. When the signal-to-noise ratio gets too high the accumulated noise can be corrected by adding the last nucleotide successfully incorporated into the elongating strand. This process is repeated whenever required, until the entire sequence is obtained. For example, if the sequence obtained using the method of the present invention is determined to be "ATGCTA", the nucleotide precursor that should be added in the subsequent trial is dTTP. Thus, if some of the primer strands had not previously been elongated up to the fifth nucleotide "T", and they are therefore two nucleotide residues shorter than the rest of the primer strands in the reaction mixture, the addition of dTTP will carry the shorter primers along so that they will always be one nucleotide shorter than the majority of the primer molecules in the population; as a result, the "mistakes" will be carried along for eliminating when necessary, i.e., when the background or noise levels are too high. The mistake can be corrected in the example above by then repeating the addition of dAPT.

In an alternate embodiment of the method for sequencing DNA of the present invention, after each incorporation of a nucleotide precursor an additional quantity of the same precursor is added to ensure that all the primer ends have been elongated to the same position. This procedure is repeated until the entire sequence is obtained.

5.4.2. SITE-SPECIFIC MUTAGENESIS

Since the sequence of the DNA or RNA template is known, the nucleotide precursors are added in the proper sequence in sufficient excess and for sufficient time to ensure complete reaction. The nucleotide precursors can be added one at a time, as in the program reproduced in the example, or up to three nucleotide precursors can be added at a time.

For example, for the synthesis illustrated in FIG. 3, the base that would normally fit in the site to be modified is a "C". To a solution of the template, primer and polymerase are added excess amounts of dTTP and dATP. After the unreacted precursors are flushed out, an excess of dCTP is added to

complete the synthesis through the third base. After the dCTP has been flushed, excess amounts of dGTP, dATP and dTTP are added to take the synthesis up to the site to be modified.

Alternatively, a system with the DNA template and primer immobilized on an insoluble inert support is employed in which the polymerase is passed through the reactor with the nucleotide precursor, and the polymerase is retained in a molecular trap below the outlet valve of the reactor. During the synthesis of the portion of the strand before the nucleotide to be modified, the polymerase is recycled by back-flushing the trap. At that point, however, an error-prone polymerase (that is, one with no proof-reading function, such as avian myeloblastoma virus polymerase) is used with one equivalent of the nucleotide precursor to be incorporated. After the error-prone polymerase is flushed from the reactor, the synthesis sequence continues with the appropriate nucleotide precursors and the recycled high-fidelity polymerase.

5.5. DETECTION OF UNINCORPORATED NUCLEOTIDE PRECURSORS IN THE EFFLUENT

Any quantitative assay for nucleotides may be used to detect the unincorporated nucleotides in the effluent. Some methods which may be used to detect unincorporated precursors in the effluent include but are not limited to spectroscopic methods, such as absorption or fluorescence spectroscopy; radioactive labeling and counting (provided the nucleotide precursors are radiolabeled); and electrochemical or conductivity methods. Some of these methods are discussed in more detail below.

5.5.1. ABSORPTION SPECTROSCOPY

Since nucleotides absorb light in the ultraviolet range (e.g. at a wavelength of 254 nm or a range of 250 to 280 nm), the amount of absorbance of the effluent in the ultraviolet range can indicate whether or not the specific nucleotide added to the reaction chamber was incorporated into the growing nucleotide chain. More specifically, absorbance of the effluent correlates inversely with incorporation of the nucleotide precursor into the growing nucleotide chain. That is, if the nucleotide precursor was incorporated into the growing nucleotide chain, then less nucleotide will be present in the effluent and the absorbance of the effluent will register at a correspondingly lower level (i.e., there will be less increase in the absorbance reading). If the nucleotide precursor was not incorporated into the growing chain, then all of the added nucleotide will be washed out with the effluent and the absorbance of the effluent will increase. In other words, the relative absorbance of the effluent indicates whether the specific nucleotide precursor added to the reaction chamber was incorporated into the chain and therefore is the next base of the nucleotide sequence. An example of the steps of this embodiment of the present invention is illustrated schematically in FIG. 2. Ideally, the elongation should proceed in a synchronous fashion.

According to one embodiment of the present invention, absorbance of the effluent can be

measured at a wavelength of 254 nm. Each nucleotide precursor has its own specific absorption maximum i.e., the wavelength of maximum absorption), however, as a result the absorbance of the effluent could be reset at the specific wavelength absorption maximum for each nucleotide precursor added when reading the effluent from that particular reaction.

In another embodiment of the present invention, the effluent can be passed through a detector cell coated on one side with an illuminated fluorescing material that is excited at a wavelength absorbed by the nucleotides. The presence of nucleotides in the effluent can thus be detected by the quenching of the fluorescence.

5.5.2. DETECTION OF LABELLED PRECURSORS

The nucleotide precursors can be labeled with an appropriate radioisotope as ^{32}P , ^3H , or ^{35}S (beta-emitters), or ^{131}I (a gamma-emitter). The effluent can be assayed for the presence of the radiolabeled precursor using any appropriate detector, preferably equipped with a flow-cell. For example, the measurement of beta-activity can be accomplished by instruments which employ gas ionization methods (e.g., a Geiger-Müller counter or scintillation counting). Similarly the measurement of gamma activity can be detected and counted by means of commercially available gamma counters.

5.5.3. ELECTROCHEMICAL DETECTION

Electrochemical detection involves applying an electric potential across a flow cell in a manner that will induce oxidations or reductions of chemicals that can undergo these reactions.

5.5.4. CONDUCTIVITY DETECTION

Conductivity detection involves measuring the resistance of a solution in a flow cell to an applied electric potential. Very slight changes in the solute concentration can cause a detectable change in the resistance of the solution.

5.6. AUTOMATION AND USE OF THE COMPUTER

Ideally the method of the present invention can be automated and regulated by a computer that is programmed to control the selection and addition of the precursors, record the absorbance reading obtained for the effluent from each added nucleotide, as well as print out the sequence thus obtained. A flow chart of such a program is represented in FIG. 4. According to the scheme presented in FIG. 4, the successful incorporation of a nucleotide is followed by the addition of the penultimate nucleotide precursor as described in Section 5.4.

The first example refers to the use of the reactor represented in FIG. 5 connected to a high pressure liquid chromatography pump and ultraviolet detector. The pump, switching valves, and injector are controlled by a dedicated computer which executes programs written in the basic language, such as the one reproduced in the second example.

6. EXAMPLE: DETERMINATION OF SEQUENCE OF AN OLIGONUCLEOTIDE

The reaction vessel was charged, through the opened load holes, with 200 μ l of M13 Mp10 DNA with hybridized 17-base primer. The vessel was sealed and washed with buffer until the OD254 readings stabilized. The first base that was tested for incorporation was of GTP. After incubation, integration of the OD254 reading of the effluent indicated that .07 nmoles of dGTP had been incorporated. An additional incubation with dGTP yielded no further incorporation. Reaction with excess dCTP and integration of the OD254 reading of the effluent denoted the incorporation of .16 nmoles of dCTP. The solution in the reaction vessel was then recovered and it was determined that approximately .11 nmoles of DNA were present.

It was concluded that the primer strand had been elongated by one G, followed by two C's; therefore the sequence of the first three nucleotides past the primer sequence of the template strand was CGG.

7. EXAMPLE: THE COMPUTER PROGRAMS USED

In the sequencing program that follows, the "SEQ" program establishes the basic operating parameters for a sequencing run. The operator is first given the choice between automatic or operator control; if the automatic option is chosen the next decision is whether or not the noise elimination route will be used and, if so, after how many bases have been processed. In an automatic run it is essential that the rotary valve initially be correctly positioned; therefore an encouraging prompt is placed on the screen for the operator to check the valve setting.

Additional operating parameters are now examined and may be changed by the operator: (1) wash flow rate; (2) load flow rate; (3) reaction flow rate (usually zero); (4) travel time to reactor; (5) reaction time; (6) threshold limits to determine base usage; (7) overall run time. These values may then be printed on hard copy if desired. The program now exits to program "W". The "W" program starts washing the system while also establishing a data file to store the sequence as it is determined. "W" then calls up the "LAT" program which is responsible for injecting, incubating, and evaluating the uptake of dATP. "LAT" first reduces the flow to the load flow rate and sets the injector to load. It then sets the rotary valve to load dATP. The analysis file "A" is then called up to help control the run and acquire the data. First the aliquot of dATP is injected and travels to the reactor where it is incubated for the time set in "SEQ". Simultaneously the rotary valve is set to wash. After incubation, the reaction vessel is washed into the monitor, and absorption peak integration and evaluation begins. The results of the run are displayed on the screen. If incorporation occurred, the base is entered into the data file; if not the next base to be tested is determined. If no successful base incorporation has yet occurred the bases are simply tried in the order A, C, G and T. If there has been a previous incorporation the next base to be tried will be the penultimate one incorporated. This allows for any incompletely

stepped chains to be carried one step behind the current successfully incorporated base. As determined in the "SEQ" program, these incomplete chains can be brought into phase by repeating the last incorporated base after trying the penultimately incorporated base.

The mutagenesis program controls the sequential addition of nucleotide precursors to the growing nucleotide chain in a manner similar to the sequencing procedure. Instead of the sequence of base addition being determined by incorporation, however, the addition of bases is regulated by the sequence input by the user. The program sequentially adds the proper bases up to the point for mutagenesis, at which point the mutagenizing base is added. After its incorporation all the bases are added and replication goes to completion.

7.1. THE SEQUENCING PROGRAM

BASIC: SEQ

```

1 PRINT "COPYRIGHT 1985 DR. ROBERT J. MELAMEDE"
2 FOR Y=1 TO 100
3 NEXT Y
5 PRINT #P,"COPYRIGHT 1985 DR. ROBERT J. MELAMEDE"
10 CLEAR
12 A=0
13 Z=0
14 QS="0"
15 PS="0"
16 TS="0"
17 INPUT "IS THIS A MANUAL INJECT TEST RUN?"$
18 IF TS="Y" GOTO 30
19 INPUT "DO YOU WISH TO ACTIVATE NOISE ELIMINATION PROCESS?"$
20 IF NS="Y" INPUT "HOW MANY BASES BEFORE NOISE ELIMINATION?"N1
30 PRINT "DATP=#1 DCTP=#2 DGTP=#3 DTTP=#4 WASH=#5 (&5)"
40 PRINT "DO YOU WISH TO STEP THE VALVE?"
50 INPUT $S
60 IF $S="N" THEN GOTO 145
70 IF $S="Y" THEN 80
80 XCOM 10FF
90 XCOM 20N
100 FOR Y=1 TO 100
110 NEXT Y
120 XCOM 20FF
130 XCOM 10N
140 GOTO 40
145 CLEAR
150 PRINT "TO CHANGE THE DEFAULT VALUES () OF CONTROL VARIABLES"
160 PRINT "ENTER Y FOLLOWED BY RETURN"
165 PRINT
170 INPUT "DO YOU WANT TO CHANGE THE WASH FLOWRATE (.5)?"Y$
180 IF Y$="Y" THEN GOTO 200
190 IF Y$="N" GOTO 220
200 INPUT "ENTER THE NEW WASH FLOWRATE" W
210 GOTO 230
220 W=.5
230 INPUT "DO YOU WANT TO CHANGE THE LOAD FLOWRATE (.2)?"Y$

```

```

5004 PRINT "WASH TIME=1"
5005 PRINT "MUTATIONAL REACTION TIME=2"
5006 PRINT "COMPLETION REACTION=10"
5010 INPUT "IF YOU WANT TO CHANGE ANY OF
THEM ENTER Y*Y$
5020 IF Y$="Y" THEN 5035
5030 GOTO 9
5035 CLEAR
5040 PRINT "TO CHANGE DEFAULT VALUES
ENTER Y & RETURN"
5050 INPUT "DO YOU WANT TO CHANGE THE
FLOW RATE (2)*Y$
5055 F=2
5066 T=1
5060 IF Y$="Y" THEN 5080
5070 GOTO 5140
5080 INPUT "ENTER NEW FLOWRATE" F
5090 PRINT "THE PRODUCT OF FLOW-
RATE*TRAVEL TIME MUST EQUAL"
5100 PRINT "THE REQUIRED LOADING VOLUME
THIS WILL AUTOMATICALLY"
5110 PRINT ".ALTERED"
5120 T=2/F
5121 PRINT "THE NEW TRAVEL TIME VALUE
IS" PRINT T
5140 INPUT "DO YOU WANT TO CHANGE REAC-
TION TIME (2)*Y$
5145 R=2
5150 IF Y$="Y" THEN 5170
5160 GOTO 5180
5170 INPUT "ENTER THE NEW REACTION TIME
VALUE" R
5180 INPUT "DO YOU WANT TO CHANGE THE
WASH TIME (1)*Y$
5190 W=2
5200 IF Y$="Y" THEN 5220
5210 GOTO 5230
5220 INPUT "ENTER THE NEW WASH TIME
VALUE" W
5230 PRINT "DO YOU WANT TO CHANGE THE
MUTATIONAL REACTION TIME"
5240 INPUT " (2)*Y$
5250 M=2
5260 IF Y$="Y" THEN 5280
5270 GOTO 5290
5280 INPUT "ENTER NEW MUTATIONAL REACTION
TIME" M
5290 PRINT "DO YOU WANT TO CHANGE THE
COMPLETION REACTION"
5300 INPUT " TIME(10)*Y$
5310 F=10
5320 IF Y$="Y" THEN 5340
5330 GOTO 5400
5340 INPUT "ENTER NEW REACTION COMPLE-
TION TIME" F
5400 PRINT #P,"FLOW",F
5410 PRINT #P,"TRAVEL TIME",T
5420 PRINT #P,"REACTION TIME",R
5430 PRINT #P,"WASH TIME",W
5440 PRINT #P,"MUTATIONAL REACTION TIME",M
5450 PRINT #P,"COMPLETION REACTION TIME",F
5460 GOTO 9

```

Claims

1. A method for determining the base sequence of nucleotides comprising:

(a) adding an activated nucleoside 5'-triphosphate precursor having a known nitrogenous base to a reaction mixture comprising a template-directed nucleotide polymerase and a single-stranded polynucleotide template hybridized to a complementary oligonucleotide primer so that the primer is at least one nucleotide residue shorter than the template at the 3'-end of the primer under reaction conditions which allow incorporation of the activated nucleoside 5'-triphosphate precursor onto the primer at the 3'-end of the primer provided the nitrogenous base of the activated nucleoside 5'-triphosphate precursor is complementary to the nitrogenous base of the unpaired nucleotide residue of the template; and

(b) detecting whether or not the nucleoside 5'-triphosphate precursor was incorporated into the primer strand.

2. The method according to claim 1 in which the detection of incorporation of the activated precursor is accomplished by:

(a) separating unincorporated nucleoside 5'-triphosphate precursors from the reaction mixture; and

(b) detecting the presence or absence of the unincorporated nucleoside 5'-triphosphate precursors in the separated component.

3. The method according to claim 2 in which the presence or absence of the separated unincorporated nucleoside 5'-triphosphate precursor is detected by absorption spectroscopy.

4. The method according to claim 3 in which the absorption wavelength is in the ultraviolet range.

5. The method according to claim 2 in which the presence or absence of the separated unincorporated nucleoside 5'-triphosphate precursor is detected by fluorescence spectroscopy.

6. The method according to claim 1 in which the detection of incorporation of the precursor is accomplished by:

(a) separating unincorporated nucleoside 5'-triphosphate precursors from the reaction mixture; and

(b) detecting the presence or absence of the separated unincorporated nucleoside 5'-triphosphate precursor electrochemically.

7. The method according to claim 2 in which the activated nucleoside 5'-triphosphate precursor is radiolabeled and the presence or absence of separated unincorporated radiolabeled nucleoside 5'-triphosphate precursors

is detected by radioactive counting.

8. The method according to claim 1 or 2 in which the template comprises DNA, the nucleotide polymerase comprises a DNA-directed DNA polymerase, and the activated nucleoside 5'-triphosphate precursors comprise deoxyribonucleoside 5'-triphosphate precursors.

9. The method according to claim 1 or 2 in which the template comprises DNA, the nucleotide polymerase comprises a DNA-directed RNA polymerase, and the activated nucleoside 5'-triphosphate precursors comprise ribonucleoside 5'-triphosphate precursors.

10. The method according to claim 1 or 2 in which the template comprises RNA, the nucleotide polymerase comprises an RNA-directed DNA polymerase and the activated nucleoside 5'-triphosphate precursors comprise deoxyribonucleoside 5'-triphosphate precursors.

11. A method for the site-specific mutagenesis of a nucleotide, comprising:

(a) adding an activated nucleoside 5'-triphosphate precursor having a known nitrogenous base to a reaction mixture comprising a template-directed nucleotide polymerase and a single-stranded polynucleotide template having a known sequence hybridized with a primer that is at least one nucleotide residue shorter than the template at the 3'-end of the primer under reaction conditions which allow incorporation of the activated nucleoside 5'-triphosphate precursor onto the primer at the 3'-end of the primer provided the nitrogenous base of the activated nucleoside 5'-triphosphate precursor is complementary to the nitrogenous base of the unpaired nucleotide residue of the template; and

(b) detecting the incorporation of the 5'-triphosphate precursors by separating any unincorporated precursor from the reaction mixture and detecting the relative amount of unincorporated nucleotides in the separated component;

(c) repeating steps 11(a) and 11(b) until the primer is elongated up to the nucleotide residue on the template preceding the residue which is to be mutagenized;

(d) adding an analog of a nucleoside 5'-triphosphate precursor having a base comprising the desired mutation which will mis-pair with the template under conditions which allow incorporation of the analog into the primer strand; and

(e) detecting the incorporation of the analog by separating any unincorporated analogs from the reaction mixture and detecting the relative amount of unincorporated analogs in the separated component; and

(f) adding activated nucleoside 5'-triphosphate precursors comprising a mixture of all four bases to the reaction mixture so that the synthesis of the

mutagenized nucleotide is completed.

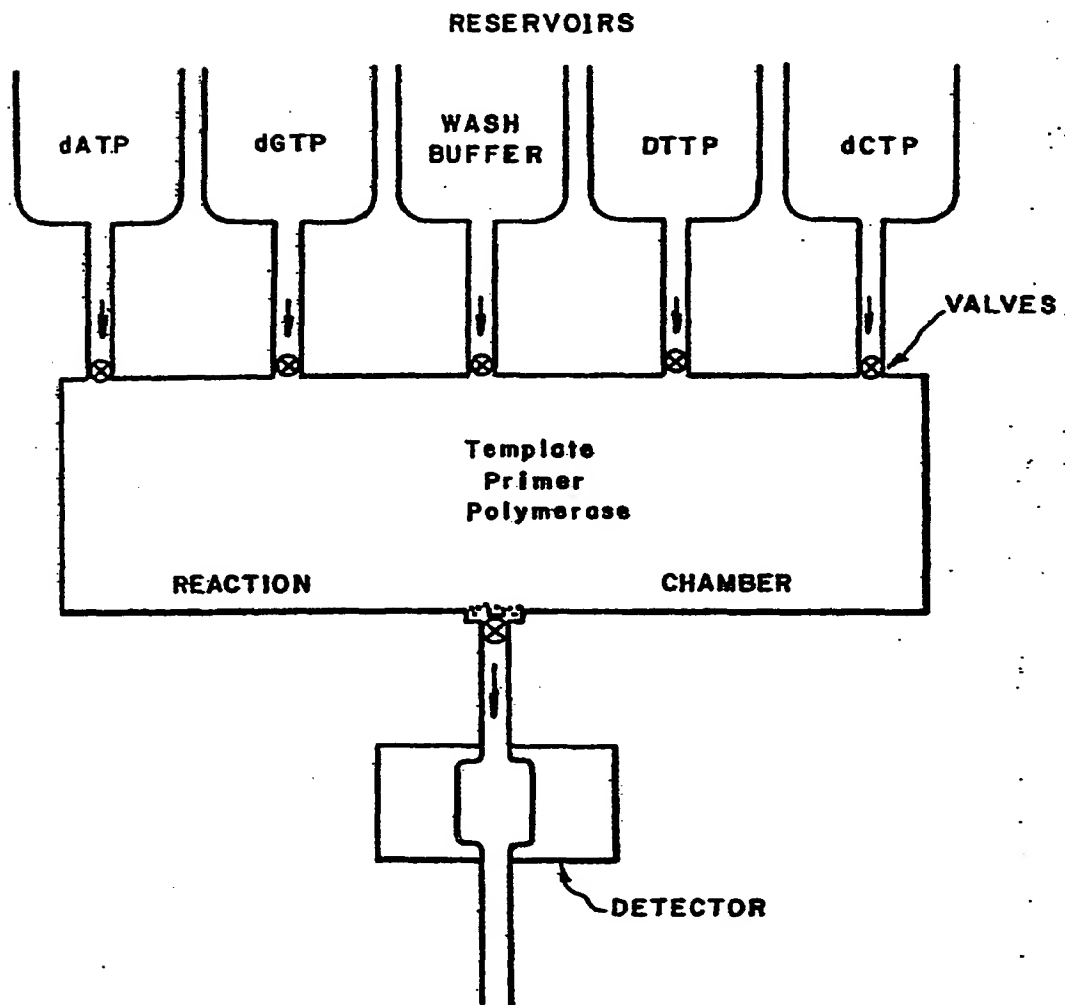
12. The method according to claim 11 in which the analog of the nucleoside 5'-triphosphate comprises a nucleoside (1-thio)-triphosphate.

0223618

Neu eingereicht / Newly filed
Nouvellement déposé

FIG. 1

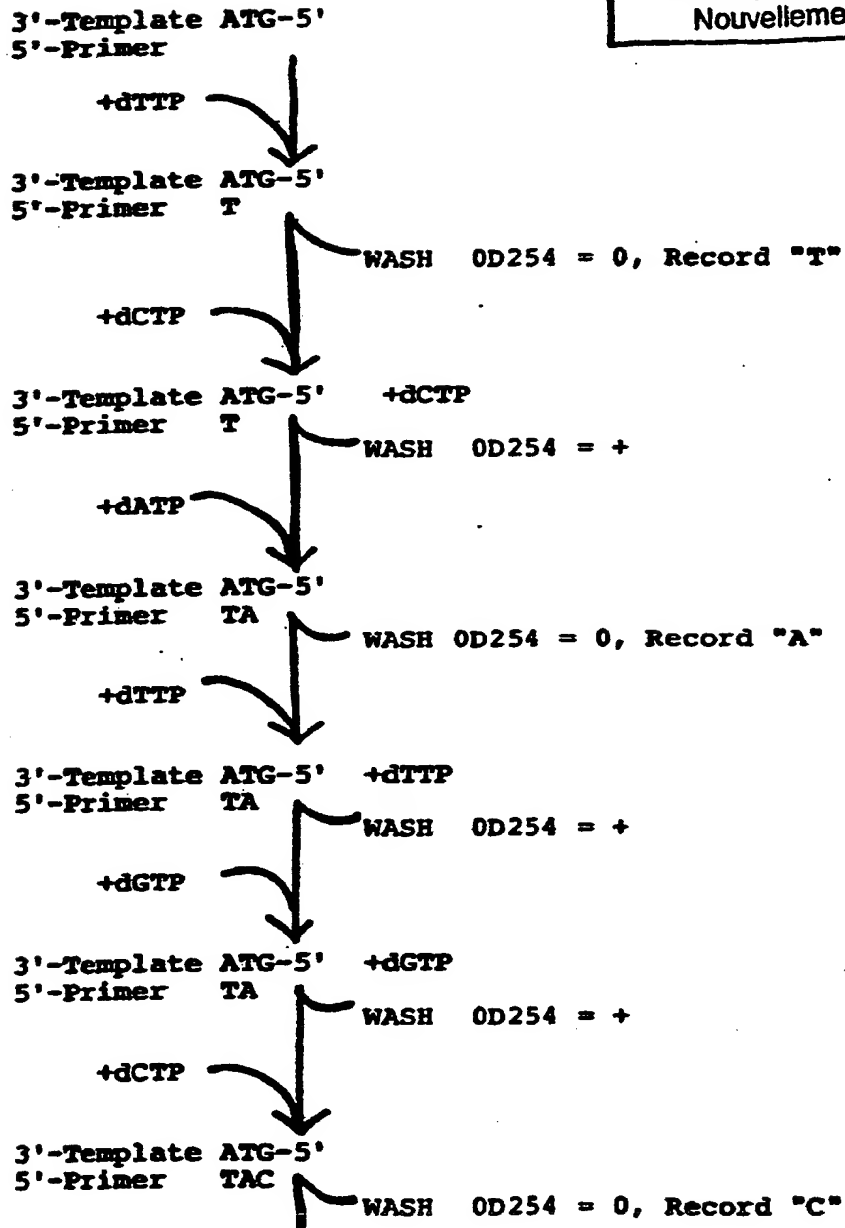
SCHEMATIC DIAGRAM OF COMPONENTS



0223618

FIGURE 2
DNA SEQUENCING

Neu eingereicht / Newly filed
Nouvellement déposé

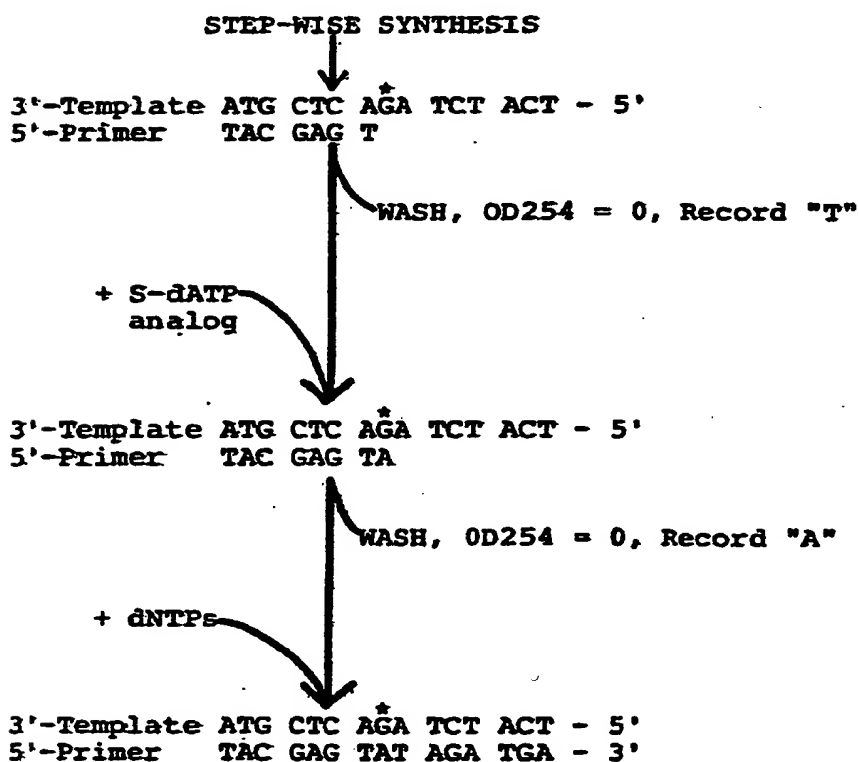


29-09-88

0223618

Neu eingereicht / Newly filed
Nouvellement déposé

FIGURE 3
SITE-SPECIFIC MUTAGENESIS



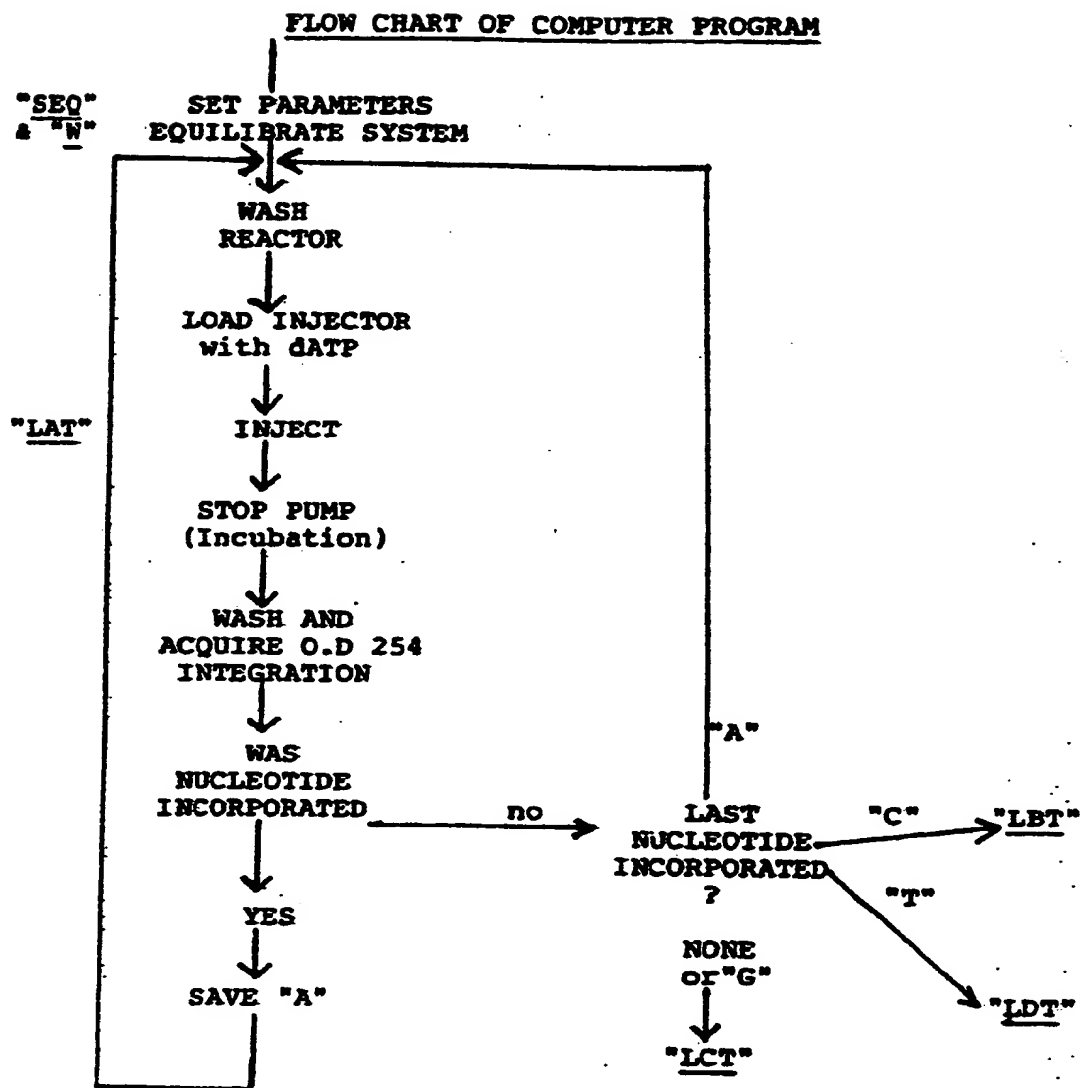
* = site of desired mutational base substitution

29.09.85

0223618

Neu eingereicht / Newly filed
Nouvellement déposé

FIGURE 4



["LBT", "LCT", and "LDT" are subprograms of identical function to "LAT", for dCTP, dGTP, and dTTP, respectively]

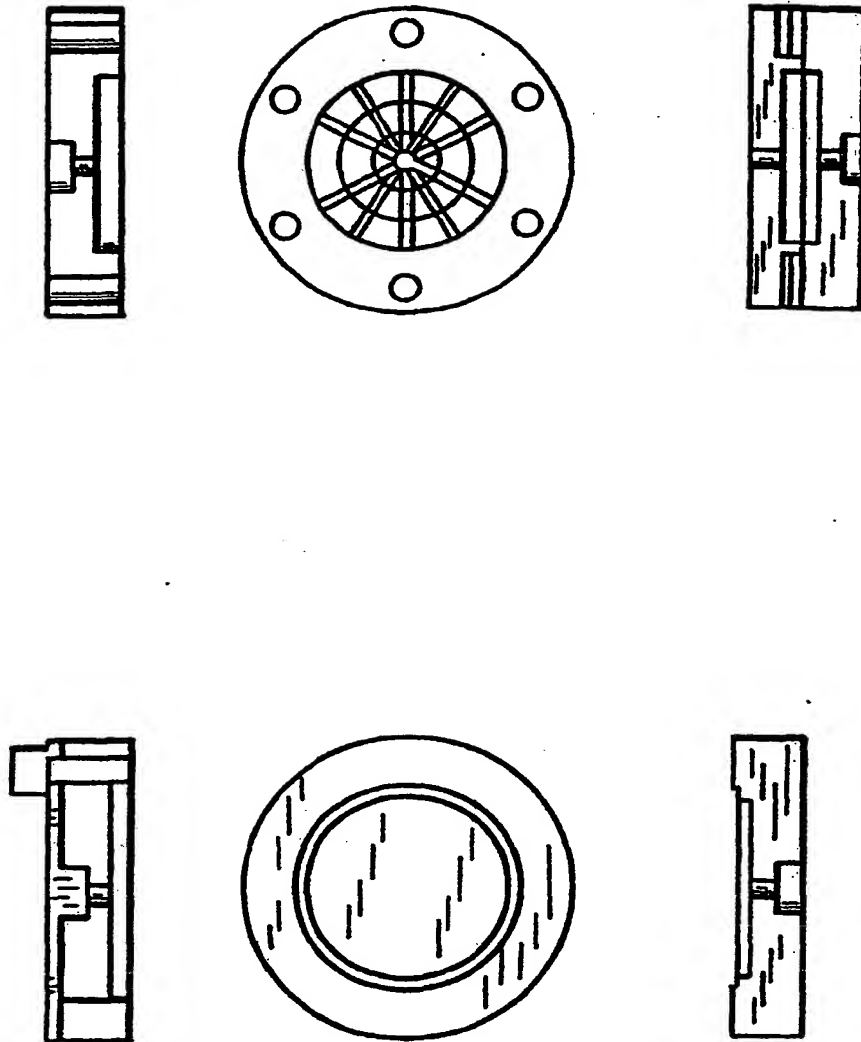
29.09.88

0223618

Neu eingereicht / Newly filed
Nouvellement déposé

DIAGRAM OF REACTION CHAMBER

FIG. 5



12

EUROPEAN PATENT APPLICATION

21 Application number: 86401590.4

91 Int. Cl.: **C 12 Q 1/68**
// G01N33/48, G01N21/33

22 Date of filing: 16.07.86

30 Priority: 18.07.85 US 756129

43 Date of publication of application:
27.05.87 Bulletin 87/22

64 Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

68 Date of deferred publication of search report:
08.07.87 Bulletin 87/28

71 Applicant: **NEW YORK MEDICAL COLLEGE**
Elmwood Hall
Valhalla New York (US)

72 Inventor: **Melamede, Robert Jean**
Dean road
Carmel New York 10512 (US)

74 Representative: **Warcoin, Jacques et al**
Cabinet Régimbeau 26, avenue Kléber
F-75116 Paris (FR)

54 Automatable process for nucleotide sequencing.

57 This invention relates to a method for determining the nucleotide sequence of DNA and RNA molecules. The method is automatable and avoids the use of radioactive labels and gel electrophoresis. The method is also adaptable for introducing site-specific mutations in DNA and RNA molecules.